04

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/770,421	02/04/2004	David Llewellyn Rees	01263.001020.1	1942

5514 7590 04/17/2006

FITZPATRICK CELLA HARPER & SCINTO
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

EXAMINER	
LERNER, MARTIN	
ART UNIT	**PAPER NUMBER**
2626	

DATE MAILED: 04/17/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/770,421	REES, DAVID LLEWELLYN	
	Examiner	Art Unit	
	Martin Lerner	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *28 March 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *13, 14, 16 to 24, 37, 38, 40 to 48, 50, and 52* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *13, 14, 16 to 24, 37, 38, 40 to 48, 50, and 52* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☒ Certified copies of the priority documents have been received in Application No. 09/409,247.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 112

1. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

2. Claims 13, 14, 16 to 24, 37, 38, 40 to 48, 50, and 52 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claims contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

The limitations of "speech detection means operable to process said received signal and to identify when speech is present in the received signal" and "wherein said likelihood determining means is operable to determine said likelihoods in the received signal when said speech detecting means detects speech within the received signal" lack enablement because Applicants' Specification does not disclose any embodiment combining speech detection means with a distinct means for determining the likelihood that said boundary is located at each of a plurality of possible locations and means for determining the location of said boundary using the likelihoods. Thus, Applicants' Specification does not enable one having ordinary skill in the art to make and/or use the invention because there is no disclosed embodiment having both a speech detection means and means for determining a likelihood of a boundary.

Applicants' means for determining the likelihood of the location of boundaries simply comprise elements of a speech detection means. Thus, it is improper to claim distinct elements relating to a speech detection means and means for determining a boundary between speech and non-speech. Applicants' Specification does not disclose that there is a distinct speech detection means for determining whether speech is present, followed by means for determining whether a boundary is present. Applicants' Specification, Pages 15 to 25, discloses a first embodiment of a speech detection means having an endpoint detector by counting frames above a threshold. Then, Applicants' Specification, Pages 25 to 29, discloses a second embodiment of determining an end point by a maximum likelihood method. Speech detection means of the first embodiment is disclosed as alternative to, but not in combination with, means for determining a likelihood of a boundary of the second embodiment. However, Applicants' Specification does not disclose any embodiment containing both speech detection means and means for determining a likelihood of a boundary that is distinct from speech detection means.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 13, 18, 21, 37, 42, 45, 50, and 52 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chigier* in view of *Gupta et al. ('055)*.

Concerning independent claims 13, 37, 50, and 52, *Chigier* discloses an apparatus, method, computer executable process, and computer executable steps, comprising:

"means for receiving the input signal" – an input speech signal 14 is received (column 4, lines 25 to 45: Figure 1);

"means for processing the received signal to generate an energy signal indicative of the local energy within the received signal" – spectral analyzer 12 performs spectral analysis (e.g., computes a short term Fourier transform) on a window of samples to provide a feature vector sequence 16, consisting of a set of parameter coefficients (e.g. cepstral coefficients) characteristic of each speech frame (column 4, lines 46 to 59: Figure 1); cepstral coefficients are "an energy signal indicative of the local energy" because they represent a log energy of a speech signal (Figures 2 and 2A);

"means for determining the likelihood that said boundary is located at each of a plurality of possible locations within said energy signal" – a boundary classifier 54 assigns to each speech frame a probability ("the likelihood") that the speech frames correspond to a boundary between two phonemes (column 6, lines 10 to 24: Figures 3 and 3A); word boundaries 44 correspond to a case in which an initial sound 50 is classified as part of background signal 52 ("background noise containing portion") (column 5, line 64 to column 6, line 9: Figures 2 and 2A);

"means for determining the location of said boundary using said likelihoods determined for each of said possible locations" – if a boundary probability assigned to a speech frame is greater than a first threshold (e.g., 70%), the frame is assumed to be a boundary by a segment generator 56, which generates a network of speech segments (A, B, and C); in operation, boundary classifier classifies boundaries I, II, and III in a speech frame sequence 59; segment generator 56 produces speech segments A, B, and C based on the classified boundaries (column 6, lines 15 to 38: Figures 3 and 3A).

Concerning independent claims 13, 37, 50, and 52, *Chigier* discloses detecting whether speech is present by classifying boundaries, but omits speech detection means distinct from means for determining a likelihood of a boundary, for limitations of "speech detection means operable to process said received signal and to identify when speech is present in the received signal" and "wherein said likelihood determining means is operable to determine said likelihoods in the received signal when said speech detecting means detects speech within the received signal." However, *Gupta et al. ('055)* teaches a voice activity detector for speech signals in variable background noise, where a voice activity detector (VAD) flag is employed to discriminate between speech and silence and adapt to background noise. Stated advantages are to detect speech with minimal clipping and false alarms. When a VAD flag is set to one, then speech is compared to a first threshold, and when a VAD flag is set to zero, then speech is compared to a second threshold. (Column 1, Line 28 to Column 2, Line 15; Column 5, Lines 1 to 64: Figures 4 to 6) It would have been obvious to one having ordinary skill in the art to incorporate a VAD flag as speech detection means taught by *Gupta et al.*

('055) in an apparatus, method, computer executable process, and computer executable steps for determining boundary likelihoods of *Chigier* for the purpose of adapting to variable background noise with minimal false alarms.

Concerning claims 18 and 42, *Chigier* discloses spectral analyzer 12 blocks a sampled speech signal into frames by placing a "window" over the samples that preserves the samples in the time interval of interest (column 4, lines 45 to 50: Figure 1A).

Concerning claims 21 and 45, *Chigier* discloses word boundaries 44 correspond to a case in which an initial sound 50 is classified as part of background signal 52 (e.g. when sound 50 is a typical mouth click or pop produced by opening the lips, prior to speaking), and boundaries 46, correspond to a case in which an initial sound is classified as part of a word (column 5, line 64 to column 6, line 9: Figures 2 and 2A); implicitly, at least a boundary at a beginning of a speech portion is detected.

5. Claims 14, 22, 38, and 46 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chigier* in view of *Gupta et al. ('055)* as applied to claims 13 and 37 above, and further in view of *Cohrs et al.*

Concerning claims 14 and 38, *Chigier* discloses checking boundary probability classifications of one or more frames from either side of frame N (column 6, line 65 to column 7, line 1), but omits determining a boundary location by comparing with a model representative of energy in background noise and a model representative of energy in

speech, and combining results of the comparisons to determine a likelihood for a current location. However, *Cohrs et al.* teaches computation of a similarity measure between stored references and parameters extracted from an utterance using hidden Markov models (HMMs). Hypothesizer 43 makes two types of hypotheses. The first type of hypothesis (referred to as a "background hypothesis") assumes that the feature vector sequence includes only background. The second type of hypothesis (referred to as a "phrase hypothesis") assumes that the feature sequence includes a command word. (Column 4, Line 59 to Column 5, Line 20: Figure 2) *Cohrs et al.* states there is an advantage in using models instead of thresholds for spotting command words by avoiding problems associated with false alarm rates for certain users. (Column 1, Lines 31 to 63) It would have been obvious to one having ordinary skill in the art to determine boundaries by comparing to models of background noise and speech as taught by *Cohrs et al.* in the method and apparatus for boundary probability assignment of *Chigier* for the purpose of avoiding problems associated with using thresholds.

Concerning claims 22 and 46, *Cohrs et al.* teaches hidden Markov models (HMMs) (column 4, lines 1 to 5), which are statistical models, implicitly.

6. Claims 16, 17, 19, 20, 40, 41, 43, and 44 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chigier* in view of *Gupta et al. ('055)* as applied to claims 13 and 37 above, and further in view of *Lennig et al.*

Concerning claims 16, 17, 40, and 41, *Chigier* omits filtering an energy signal to remove energy variations having a frequency below a predetermined frequency, where

the filter is operable to filter out energy variations below 1 Hz. However, *Lennig et al.* teaches detecting word endpoints, where filter means 12 comprises a filter bank of twenty triangular filters spanning a range of about 100 Hz to about 4000 Hz. Weights W_{ij} for filter channels j are set so that $W_{ij} = 0$ for frequencies l_j below 100 Hz. (Column 3, Lines 4 to 40: Figure 1; Table 1: Filter No. 1) Thus, all energy variations at frequencies in the range between 0 Hz and 100 Hz are removed, including those energy variations at frequencies below 1 Hz. *Lennig et al.* suggests an advantage of reducing an error rate for speech recognition. (Column 1, Lines 19 to 26) It would have been obvious to one having ordinary skill in the art to filter an energy signal to remove energy variations having a frequency below a predetermined frequency as taught by *Lennig et al.* in the method and apparatus of boundary probability assignment of *Chigier* for the purpose of reducing an error rate for speech recognition.

Concerning claims 19, 20, 43, and 44, *Chigier* discloses speech samples (column 4, lines 60 to 66), and assigning boundary probabilities based on log energy (column 6, lines 10 to 24: Figures 2, 2A, and 3).

7. Claims 23 and 47 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chigier* in view of *Gupta et al. ('055)* and *Cohrs et al.* as applied to claims 13, 14, 22, 37, 38, and 46 above, and further in view of *Abut et al.*

Cohrs et al. discloses hidden Markov models (HMMs), but omits models based on Laplacian statistics. However, *Abut et al.* discloses speech probability models based on Laplacian speech statistics. (II. Speech Statistics: Page 226) It is suggested that

Laplacian statistics have lower and upper bounds suitable for speech probability models. (Page 227) It would have been obvious to one having ordinary skill in the art to utilize models based upon Laplacian statistics as suggested by *Abut et al.* in the method and apparatus for boundary probability assignment of *Chigier* in order to obtain suitable speech probability models.

Claims 24 and 48 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Chigier* in view of *Gupta et al. ('055)* and *Cohrs et al.* as applied to claims 13, 14, 22, 37, 38, and 46 above, and further in view of *Erell et al.*

Cohrs et al. discloses hidden Markov models (HMMs), but does not expressly state that a speech model is an auto-regressive model. However, *Erell et al.* teaches a speech recognition system where the acoustic features are extracted to form a feature vector, and where the features are the coefficients of an autoregressive model. *Erell et al.* states that these are the most commonly used features, including linear prediction coefficients, cepstrum coefficients, bank of filter energies etc., to reflect vocal tract characteristics. (Column 1, Lines 37 to 45) It would have been obvious to one of ordinary skill in the art to use an auto-regressive model in the method and apparatus for boundary probability assignment of *Chigier* because *Erell et al.* suggests that an auto-regressive model is the most commonly employed method of deriving speech features.

Response to Arguments

8. Applicants' arguments filed 28 March 2006 have been considered but are moot in view of the new grounds of rejection.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ML
4/13/06

Martin Lerner
Examiner
Group Art Unit 2626